

交通网络的通行时间预测与最优路径决策

摘要:

本文首先分析了道路系统的微观物理模型，由微观模型推导出了宏观模型并得出了一些重要结论，进一步在此基础上建立了一维交通线路的时间预测问题的三种算法并讨论了这些算法对测量数据的要求。通过理论论述和 TransGuide 系统的实测数据验证了模型的正确性和时间预测算法的有效性。然后，在对交通网络的宏观统计特征的分析基础上，分别研究了稳态独立随机交通网络与时变相关随机交通网络的时间预测与最优路径决策问题。根据人群选择的主体差异性构造了最优路径的统一的优化目标函数并进一步给出了两种有效的随机网络最优路径优化算法。最后，给出了赛题中的具体算例的最优路径。

关键词：随机网络；交通模型；时间预测；最优路径

一、问题的提出

城市的交通状况直接影响城市的竞争力、城市形象以及居民的生活的便利程度。随着城市化的进展和汽车的普及，交通拥挤加剧，交通事故频发，交通环境恶化，这成为长期以来困扰发展中国家和发达国家的严重问题。随着现代科学技术的发展智能交通系统 ITS (Intelligent Transport System) 应运而生，缓和了道路堵塞、减少交通事故，方便了使用者。它通过传播实时的交通信息使出行者对即将面对的交通状况有足够的了解并据此做出正确的选择；缓解了道路堵塞、减轻了对环境的污染、提高了行驶安全、减少了行驶时间。

在美国有两例城市道路智能交通系统的范例^[1]：一是洛杉矶圣莫尼卡市的 Smart Corridor；一是德克萨斯州圣安东尼奥的 TransGuide。事实证明，该系统的应用不仅使事故率下降了 15%，而且，根据德克萨斯州运输研究所的对比研究，运行期间的事故反应时间也缩短了 20%。在预测出行时间方面，TransGuide 的成绩也得到了 85% 的司机的肯定^[2]。

目前，ITS 交通诱导系统已经成为国际研究热点之一^{[2]~[8]}，建立了多维时间序列分析、模糊神经网络、ARMA 模型、能量谱模型、混沌小波模型等等多种理论模型。在交通流的特征、建模、时间预测以及旅行时间最短路径的选择方面起到重要的推动作用。

然而，目前的 TransGuide 系统依旧不是很完善，今年的全国研究生数学建模比赛提出了一系列问题，本文将对交通流进行理论建模，并试图通过此理论模型来完成最短时间路径优化和时间预测等问题并对赛题中的问题进行逐一回答。

二、公路微观交通模型

2.1 基本变量

- 1) 平均车速 v ：指某一截面在一时间段内通过的车辆的平均速度；
- 2) 反应时间 T ：包括司机的生理反应时间和机械操作的滞后时间；
- 3) 车流量 q ：单位时间通过某截面的车辆数量；
- 4) 车间距 l ：前后两车的中心距离；
- 5) 车辆平均长度 l_c ：公路上的大量车辆的平均长度；
- 6) 停车距 l^* ：前后车完全静止时为了安全和再次启动需要保持的净间距；
- 7) 有效车间距 l_e ：可以由来提供刹车的间距，等于 $l - l_c - l^*$ ；
- 8) 极限车间距 l_0 ：在给定的车速下，安全行车的最小车间距；
- 9) 区段长度 L ：该区段的长度；
- 10) 车数 n ：在某区段上的车子总数；
- 11) 公路的极限运输能力 $p(x)$ ：公路某截面的极限通行能力；

- 12) 车流密度 ρ : 单位长度路面上的车辆数量, 显然有 $\rho = \frac{1}{l}$ 。
- 13) 路面有效长度 L_e : 这里的净长度路面是指除去该路段上所有车辆占据的长度 nl_c 和不能由车辆支配的全部保留停车距 nl^* 当车辆全部堵住时, 显然有 $L_e = 0$;
- 14) 车流有效密度 ρ_e : 单位有效长度上的车辆数, 显然有 $\rho_e = \frac{1}{l_e}$ 。当车完全堵死时, 有 ρ_e 趋向无穷大。
- 15) 车流临界密度 $\rho_c(x)$: 能使得车流保持平均速度为 v^* 的最大车流密度。显然有 $\rho_c = \frac{p(x)}{v^*}$;

2.2 基本假设

- 1) 实际上公路车辆包含货车、客车、小汽车等多种类型, 为了问题的简化, 假定所有车辆的性能都一样;
- 2) 车流中的所有车主动以时间最短方式前进。即车流中的所有车辆在速度没有达到公路所限制的最大速度 v^* 而与前车的间距大于其行驶速度下的极限车间距时, 其会加速行驶。当速度达到 v^* 则按照 v^* 行驶。
- 3) 由于是高速公路的主干道, 所以不考虑红灯等需要暂停车流的环节, 认为车流始终是连续的, 即使堵车也不例外。
- 4) 为了问题的简化, 仅仅考虑单车道模型, 不考虑超车。所以后车对前车没有影响。
- 5) 堵车: 车流时断时续, 车流的密度达到 ρ^* ; 车流的有效密度趋向于无穷大; 车辆的有效间距为零; 车辆间距 l 等于停车距 l^* 。

2.3 跟车模型

在高速公路上, 当汽车比较多时, 后面的车不能随意地选择自己的速度, 其速度受到前面车的速度的制约。为了保证行车安全, 后车必须按照一定的方法来进行控制以获得合理的车间距^[2]。设前车的位置为 $x_n(t)$, 后车位置为 $x_{n+1}(t)$, 则车间距为

$$l(t) = x_n(t) - x_{n+1}(t) \quad (1)$$

由于前后车的速度差很小, 由假设 1 可知前后车刹车距离相同、在后车司机的反应时间内后车行驶的距离为

$$d_1 = T \cdot \dot{x}_{n+1}(t) \quad (2)$$

则有:

$$l(t) = d_1 + l^* = T\dot{x}_{n+1}(t) + l^* \quad (3)$$

其中 l^* 为停车后应该保持的距离。将 (1) 带入 (3) 则有

$$x_n(t) - x_{n+1}(t) = T\dot{x}_{n+1}(t) + l^* \quad (4)$$

两边同时取导数有:

$$\dot{x}_n(t) - \dot{x}_{n+1}(t) = T\ddot{x}_{n+1}(t) \quad (5)$$

变形得

$$\ddot{x}_{n+1}(t) = \frac{1}{T}[\dot{x}_n(t) - \dot{x}_{n+1}(t)] \quad (6)$$

式 (3) 描述了最小车间距为：

$$l_0 = d_1 + l^* = T\dot{x}_{n+1}(t) + L = v(t)T + l^* \quad (7)$$

如果 l^* 取得比较小，将近似有

$$l_0 = d_1 = Tv(t) \quad (8)$$

由 (8) 可以看出，最小车间距为车速的线性函数。由式 (6) 可以看出后车的反应为：其按照前后车的速度差的线性函数来加、减速度。一般称 $\frac{1}{T}$ 为“灵敏度”而称 $\dot{x}_n(t) - \dot{x}_{n+1}(t)$ 为前车对后车的“刺激”。

进一步，根据参数的上述定义有：

$$\rho = \frac{n}{L} = \frac{1}{l} \quad (9)$$

$$q = \frac{n}{t} = \frac{n}{L} \frac{L}{t} = \frac{1}{l} v = \rho v \quad (10)$$

2.3.1 公路非满载时的情况 当车流密度较小时，后车与前车的平均间距远远超过了最小车间距，所以后车在选择速度的时候不会受到前车的影响。可以假定进入公路的车辆成泊松分布，车辆可以达到公路所允许的上限速度 v^* 前进，所有路段的行车速度近似为常数而不受车流密度的影响。

$$q = v^* \rho \quad (11)$$

此时，速度基本为恒值而流量与车流密度成正比。

2.3.2 公路满载时的情况 当车流量继续增加，达到路的通行能力限制时，车的平均间距接近最小车间距 l_0 ，密度达到临界密度 $\rho_c(x)$ 时，后车的速度直接受到前车影响。司机被迫按照最小车间距 l_0 来行驶，由公式 (8)、(10) 得

$$q = \frac{1}{T} \quad (12)$$

此时，速度与密度完全相关，相互成反比关系，而流量基本稳定。可见公路的极限承载能力与行车速度、车流密度没有关系，而与跟车模型中反应灵敏度成线性关系。此性质与连续流体的性质很类似。所以很多人类比连续流体，采用流体力学的方法来处理中等流量的车流。

2.3.3 公路流量过载时的情况 如果涌入公路的车流量进一步增加，超过了高速公路的承载能力，必然导致部分车辆滞留从而进一步地增加了车流的密度。由于此时行车速度完全受控于车流密度，导致司机必须进一步降低速度来满足安全行车的最小车间距 $l_0(v)$ 。

而降低 $l_0(v)$ 的结果将导致车的密度进一步增大，此时，密度越来越大，车速越来越低。如果后续流量小于公路的极限通行能力，则情况会逐渐改善，如果后续依旧有超过公路极限运输能力的流量涌入公路，车间距的进一步降低，当车间距降低为停车距——有效车间距降低到零时，速度降低到了 v_* 。此时对应的车流密度趋近车流的极限密度 ρ^* 而其有效密度趋向于无穷大。此时，车流时断时续，但是在一定时间内的平均意义下的平均速度为 v_* 而非为零，同样在一定时间内的平均意义下的流量 $q(x_i, t)$ 依然为公路的极限通行能力而非为零。

由以上分析可知：短期的过载将导致平均车速的降低，从而使得该路段的通行时间的增加。但是公路的极限运输能力是由灵敏度决定的，与车流密度、平均速度无关。

如果一段时间后流量低于公路的运输能力，则车流密度开始逐渐减小，平均速度开始逐渐变大。如果持续一段时间严重过载，则当密度趋向于极限密度时，从一点开始形成堵车，逐渐延伸为一条时断时续的车队。

三、一维交通的通行时间预报

3.1 定常模型

如果一段道路，输入车流与输出车流为不随时间变化的常数且不大于公路的极限运输能力的话，公路上的交通状况（车流密度、车间距等）由道路状况唯一确定。如果道路状态不随时间变化，则交通状态不随时间变化，称之为二维定常交通流模型。此时，由于任何点处的车流密度都没有达到临界密度，所以都可以畅通无阻地按照 v^* 前进。此时通过第 i 区段的时间为

$$t_i = \frac{L_i}{v^*} \quad (13)$$

$$t = \sum_{i=1}^n t_i = \sum_{i=1}^n \frac{L_i}{v^*} = \frac{L}{v^*} \quad (14)$$

3.2 准定常模型

实际上，每段道路入口的车流 $q(x, t)$ 是随机的，而道路的通行能力也是不完全一样的，有的地方通行能力较弱形成瓶颈 $p(x_i)$ 。对于城市交通干道，在大部分时段内随机的入口流量会不可避免。如果 $q(x, t)$ 在 $p(x_i)$ 上下浮动且其均值不大于 $p(x_i)$ 。则在瓶颈点 x_i 之前，密度可能超过临界密度 ρ_c ，速度不再为 v^* 。由于交通系统是缓变的过程，所以在较小的一段预测时间内不会发生较显著的变化。从而速度仅仅是位置的函数。在每个区段 l_i 内，可以用区段两端点的速度的均值来作为该段区间的平均速度，则有：

$$v_i = \frac{v_i + v_{i+1}}{2} \quad (15)$$

$$t_i = \frac{L_i}{v_i} \quad (16)$$

$$t = \sum t_i \quad (17)$$

显然，前面的定常模型是该模型的特例。当车流量适当的时候，既形成了较为连续的车流又没有形成严重堵塞的时候可以近似为准定常模型。由于一般情况下在较小的时间段内交通状况不可能发生剧烈变化，所以在较短的时间内使用定常模型来作为对实际交通模型的近似误差也是可以接受的。

该模型能在比较小的时间段内对到达时间进行预测。但是该模型有个致命的缺点：实际情况不可能完全符合所有的计算假设，实际的交通系统受到天气、驾驶人员心理等多种因素的影响，充满高度的不确定性。当有一个区段发生严重堵塞的时候，前面的车一步一挪，后面的车的运动会延迟。这样可能导致堵车车队的后端在一段时间内是停止的。于是由传感器测量得到的平均速度可能为零。从而导致预测时间为无穷大！这显然是不合理的。有必要对此进行修正。实际上，由前面的公路过载时的特性可知：由于瓶颈点前的流量等于该段道路的极限运输能力 $p(x_i)$ ，而堵车车队的密度为极限密度 ρ^* 保持不变，从而得出堵车车队任意截面处在一段显著的时间上的平均流量依然等于 $p(x_i)$ 而不为零！这是容易引起误解的地方。所以堵车区段在较长时间上的速度并不为零，堵车车队的平均速度为：

$$v_*(x_i) = \frac{p(x_i)}{\rho^*} \quad (18)$$

显然， $v_*(x_i)$ 是个很小的数值，其与堵车前的瓶颈流量 $p(x_i)$ 成正比。所以， $p(x_i)$ 是一个重要的交通参数！

由于堵塞点下游的车流平均速度很小，所以堵塞长度的长短显著地影响通过时间。前面的理想模型假设各车辆都是尽可能地快开的。实际上，尽可能地快开并不能保证一定减少旅行时间：当前面已经形成堵车车队，则后面的车绝对不会紧紧地追尾然后来个紧急刹车。实际上，堵车是随着车流密度的增加逐渐形成的。所以车流密度的增加以及平均速度的降低以及各个区段的车流量都给出了堵车的显著信息。在堵车车队形成中，后面的车辆依次进入队尾，并按照前面所说的线性跟车原则进行控制。所以实际上，堵车车队的尾部很长一段距离内的速度都将受到其一定影响。如果详细地求解需要求解一个高维的线性常系数微分方程组。为了问题的简化我们不讨论此影响，直接考虑车辆按照前面的简化假设进入堵车队尾。

在没有突发事件的前提下，道路的通行能力 $p(x)$ 基本与时间无关，突发事件的情况我们在后面再讨论。所以堵车车队的队首必然在瓶颈点处，堵车车队的队尾必然在 $\rho = \rho^*$ 的临界点上。一般情况下，该临界点在区段的中间而非恰好在端点上。但是由于

后车受控于前车，堵车的信息会向下游传递。我们可以由区段端点处的车流密度、平均速度等信息获得堵车车队的情况。实际上，有经验的司机可以根据他附近的交通状况来判断前面是否发生堵车以及堵车的严重程度，这也体现了这个原理。

堵车长度越长则队尾越接近下游端点。那么该端点的平均速度越慢。如果该端点的速度为 $v_*(x_i)$ ，则说明堵车车队的尾部已经达到甚至超过了该端点。

实际上，堵车并没有严格的分界线，而传感器布置点都是选择在一些具有突变性的位置，所以在堵车时，区段内的速度的分布是比较接近的。不同位置的平均速度差别不会很大。所以依然可以采用前面介绍的准定常模型来解决。如果数据中，在某个时间段内其平均速度为零，则由其前面一区段的流量为瓶颈流量，计算出该堵车车队的平均速度 $v_*(x_{i+1})$ 作为该区段的前端点平均速度，后端点的平均速度依旧用实测值，计算出区段内的平均速度

$$\bar{v}_i = \frac{v_*(x_{i+1}) + v(x_i)}{2} \quad (19)$$

从而计算出该区段的平均通行时间

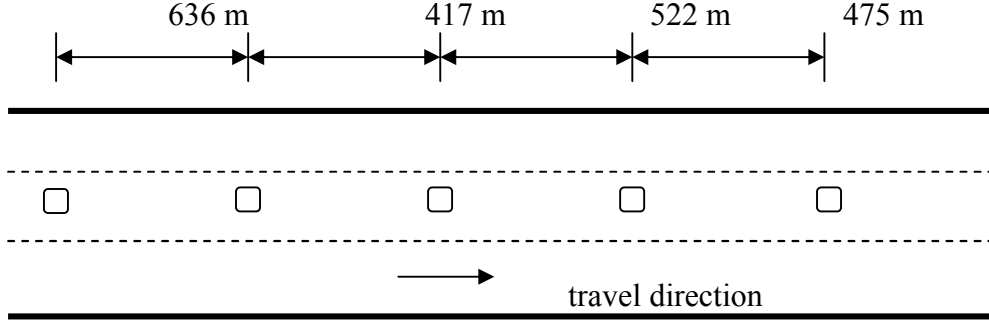
$$t_i = \frac{L_i}{\bar{v}_i} \quad (20)$$

在有大量的测量数据的情况下也可以由数据拟合出堵车长度与下游端点平均车速、区段长度之间的关系表达式，从而可以进一步较准确地预测堵车长度。

3.3 时变模型

实际上，交通系统是准周期性和复杂随机性的显著的时变系统，当需要预测的时间比较长的时候，由于系统的参数已经发生了较显著的变化，使用准定常模型就会带来较大的误差。为此考虑将更多的特征信息考虑进来，建立时变的交通模型。假设一个区段内，该区段的车流量为 $q(x,t)$ 。我们考虑连续的路段的允许流量为 $p(x,t)$ 。 $q(x,t)$ 的时变特性比较明显故将其作为随机变量。而 $p(x)$ 在相当长的一段时间内一般不会随意变化（当道路故障、车祸等突发事件发生时才会改变 $p(x)$ ）。当任意位置任意时刻 $q(x,t) \leq p(x)$ ，则不会形成交通堵塞，交通流仍近似服从定常条件。而当 $q(x_i,t) > p(x_i)$ 时，将导致部分车流受阻，使得该路段后密度开始上升，速度开始下降。当速度下降到极限速度，密度上升到饱和密度，则速度进一步将为零，发生严重堵车。

现考虑非定常条件下的时间预测问题。为了算法的实现，将连续的公路离散成 i 段，在每段内用平均密度、平均速度、平均流量来进行刻画。示意图如下图：



Detector 1 Detector 2 Detector 3 Detector 4 Detector 5

则从 t_0 到 t_1 时段内，各路段堵塞的车辆为：

$$Q(x_i, t) = \int_{t_0}^{t_1} [q(x_i, t) - p(x_i)] dt \quad (21)$$

设道路的极限密度为 ρ^* ，如果该路段的车辆密度还没有达到极限密度即 $\rho(x_i, t_1) < \rho^*$ ，那么堵塞的车辆将在整个区段内平均。则 t_1 时刻有：

$$\rho(x_i, t_1) = \rho(x_i, t_0) + \frac{1}{L_i} \int_{t_0}^{t_1} [q(x_i, t) - p(x_i)] dt \quad (\rho \leq \rho^*) \quad (22)$$

如果 $\rho(x_i, t_1) \geq \rho^*$ ，则说明该路段已经饱和，不能继续容纳堵塞的车辆，则多余的车辆向其下游段淤积。则有：

$$\rho(x_i, t_1) = \rho^* \quad (23)$$

$$\rho(x_{i-1}, t_1) = \rho(x_{i-1}, t_0) + (\rho(x_i, t_1) - \rho^*) L_i + \frac{1}{L_{i-1}} \int_{t_0}^{t_1} [q(x_{i-1}, t) - p(x_{i-1})] dt \quad (24)$$

式中的第一项为 $i-1$ 段原先的密度，第二项为 i 段多余的堵塞车辆，第三项为 $i-1$ 段的堵塞车辆。

又因为 $q(x, t)$ 只受输入的控制，受 v 、 ρ 的影响可以忽略，故除堵塞点 x_i 后的相邻区段都依然有车流。并且保持原来的车流不变。

$$q_i(t_1) = q_i(t_0) \quad (25)$$

在堵塞点后的相邻段内车流量为堵车点的瓶颈流量，即：

$$q(x_{i+1}, t_1) = p(x_i) \quad (26)$$

根据上述式子计算出

$$v_i(t_1) = \frac{q(x_i, t_1)}{\rho(x_i, t_1)} \quad (27)$$

由式 (21) ~ (23) 可以计算出经过 Δt 时刻后被预测车辆的新位置，以及各区段的新的流量、密度、平均车速。重复上述步骤进行迭代计算则可以计算出被预测车辆一直运动到目的地所需要的时间步从而得到预测时间。

3.4 突发事件的预报

在上述算法中，假设了各区段的流量是不变的。也就是说，道路的通行能力是不随时间变化的。实际问题中，往往发生一些突发事件从而导致了交通流的堵塞与疏散。

突发事件发生后，一般事发地点的通行能力会降低。假设事发后的通行能力为 $q'(x)$ ，则将 $q'(x)$ 代替 $q(x)$ 即可。再考虑突发事件处理好后，通行能力恢复正常，则用 $q(x)$ 代替 $q'(x)$ 即可。进一步， $q'(x)$ 可以由该段的流量数据获得。

当由监测数据发现某路段已经进入了过载，而其对应的流量却明显小于过去时间的流量，则由前面的分析知：必然在该区段的附近发生了突发事件从而降级了该段道路的极限运输能力。

3.5 模型的讨论

公路交通模型的研究得到了世界各国的研究者的高度重视，针对不同情景建立了各种宏观的或者微观的交通模型。一般在车流量较小时用离散模型来描述，假设进入公路的车辆符合泊松分布或者二项分布。对于连续的交通流，一般认为速度与密度成线性关系，而流量则与速度、密度成抛物线关系，当速度、密度取一合适的值 v_m 、 k_m 时，流量取得最大值。

$$v = v_f \left(1 - \frac{k}{k_j}\right) \quad (28)$$

$$q = kv_f \left(1 - \frac{k}{k_j}\right) \quad (29)$$

$$q = k_j \left(v - \frac{v^2}{v_f}\right) \quad (30)$$

式中 q 表示流量， k 表示密度， v 表示平均速度， v_f 表示最大速度， k_j 表示极限密度，其大致图象如图 1 所示。

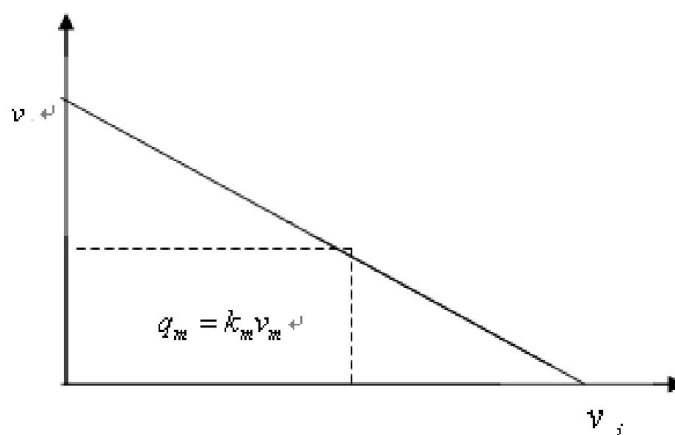


图 1 平均速度与车流密度关系图

3.5.1. 实际数据与模型之间的比较 德克萨斯州圣安东尼奥的 TransGuide 系统提供的一段 200 分钟的数据中实际的速度、流量、密度的关系分别见下图所示：

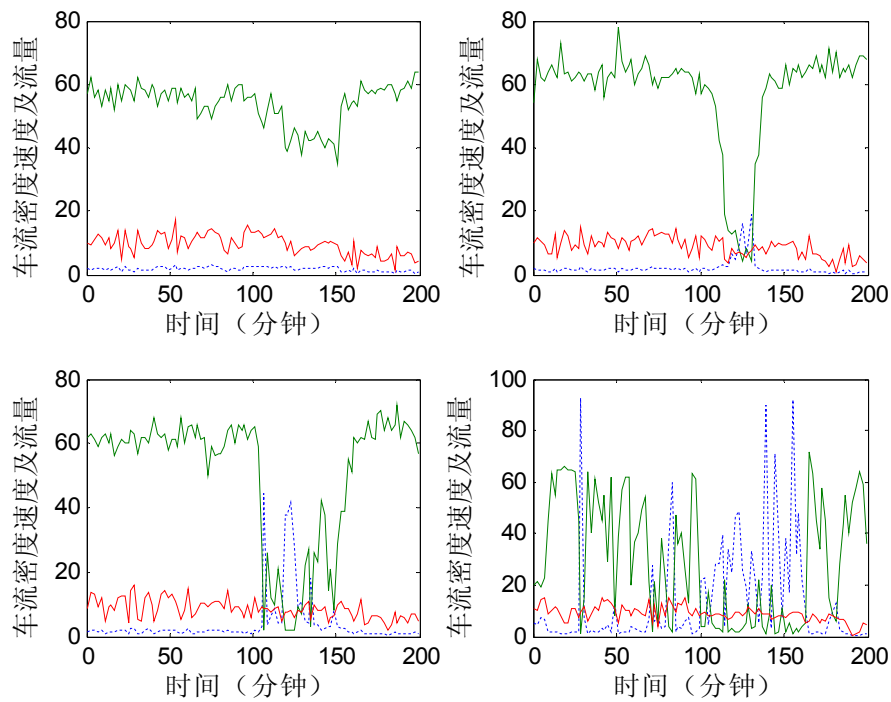


图 2 平均速度（最上端的实线）、流量（中间的实线）密度（最下段的虚线）与时间的关系

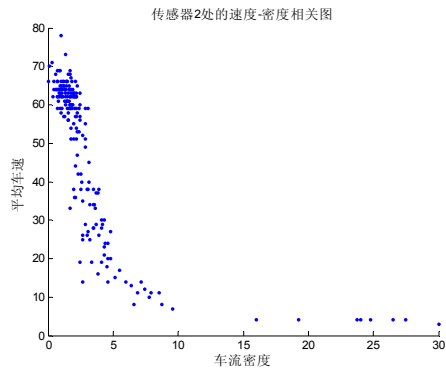


图 3 传感器 2 处的速度-密度图

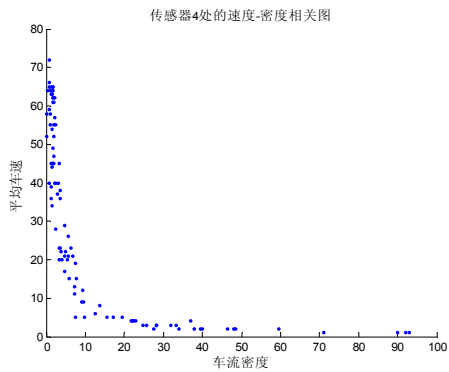


图 5 传感器 4 处的速度-密度图

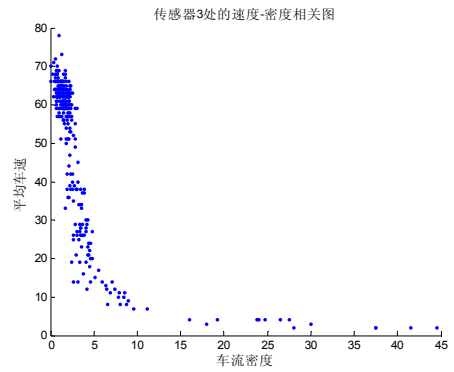


图 4 传感器 3 处的速度-密度图

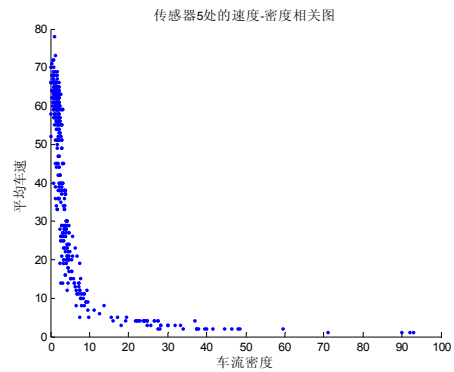


图 6 传感器 5 处的速度-密度图

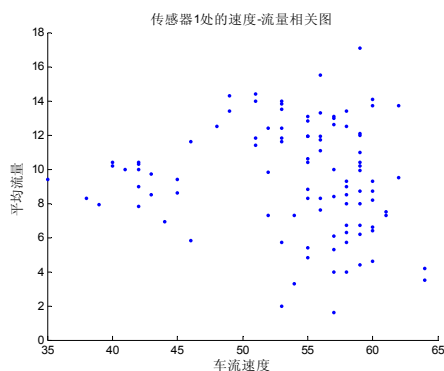


图 7 传感器 1 处的速度-流量图

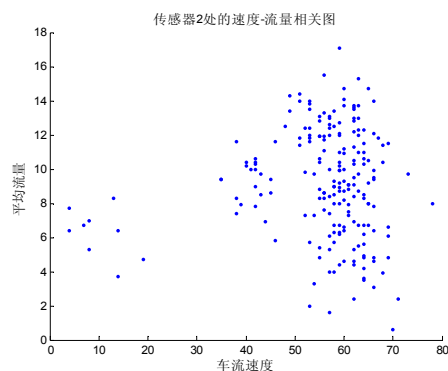


图 8 传感器 2 处的速度-流量图

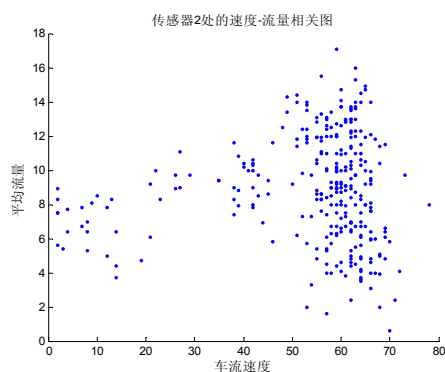


图 9 传感器 3 处的速度-流量图

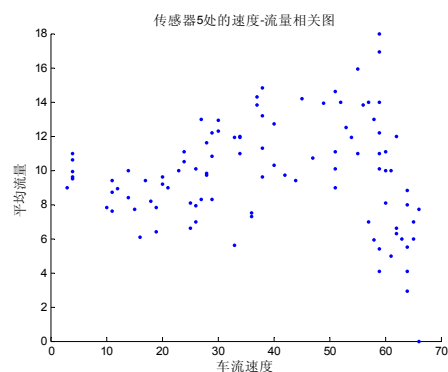


图 10 传感器 4 处的速度-流量图

从图 7~10 中可以看出：该段测量数据中，流量与速度没有显著的关系，表明流量是随机分布的，而速度并不由流量控制。由图 3~6 可以看出速度和车流密度近似成双曲线关系。而由图 2 也同样可以看出，流量与速度之间并没有显著的关系，但是平均车速与车流密度却有显著的关系。当速度小于 40 英里/小时，特别是小于 20 英里/小时，车流密度迅速提高。这点在图 3~6 中也分明看出来。

所以，该数据与笔者前面建立的准定常模型很好地吻合，而不符合一些文献中提出的线性关系。笔者建立的模型物理背景简单而清晰，对轻载、满载、超载的公路交通状况都能适应，计算量小，需要的数据量小。

3.5.2.数据采样变化时的讨论 赛题中提到了将获取数据的时间由每 2 分钟的后 20 秒变为每 20 秒钟一次的连续的采样时对模型的影响问题。对于定常模型与准定常模型，由于算法中使用的数据仅仅是速度信息，且不考虑速度的变化，所以对数据的采样频率要求不高。当获得的数据由每 2 分钟的最后 20 秒变为连续的 20 秒一次的采样，对该模型没有太多的影响。

而对于时变模型，由于考虑了车流的堵塞与疏散的时候使用了积分运算，所以必然要求时间是连续的，所以要求数据必须为连续的采样，对于每 2 分钟的后 20 秒的数据，该算法无法适用。

四、交通网络的通行时间预测与最优路径

4.1 稳态独立随机交通网络模型

实际的交通线路纵横交错，形成了双向图。交叉口为图的节点，每个两个交叉口之间的路段为图的边。类似前面讨论的一维定常模型，现在将交通网络简化为一个定常的交通网络系统。同理，在较小的时间内或者在交通比较通畅的时段内，这样的简化是合理的。

现在讨论从节点 i 到节点 j 的最优路径问题。设通过每条边 $l_{i,j}$ 的时间为相互独立的随机变量 $t_{i,j}$ ，其均值为 $\mu_{i,j}$ ，均方差为 $\sigma_{i,j}$ 。则该交通网抽象成了带有随机权重的加权有向图。设从节点 i 到节点 j 有 m 种不同的路径，其中第 k 条路径需要的通过时间为 t_k 。则 t_k 等于其通过的所有边的通行时间 $t_{i,j}$ 的和， t_k 、 $t_{i,j}$ 都为随机变量。影响 $t_{i,j}$ 的因素有很多，由中心极限定理，可以假设 $t_{i,j}$ 符合正态分布。

$$t_k = \sum t_{i,j} \quad (31)$$

由概率论知识知 t_k 也是符合正态分布的随机变量。其均值与均方差分别为

$$\mu_k = \sum \mu_{i,j} \quad (32)$$

$$\sigma_k = \sum \sigma_{i,j} \quad (33)$$

4.1.1 优化准则的确立与人群的选择偏向性 解决优化问题的第一步就是建立优化的准则。对于权值为确定常数的加权有向图，其优化的明确的准则就是路程最短。但是对于权重为随机变量的随机网络，不同的路径的通行时间同样为随机变量，不再可以直接比较大小的优劣，所以需要建立一个优化的目标函数。

由于优化是为人服务的，不同的人的选择具有显著的主体差异性。举个简单的例子：如果两条路径 C_1 与 C_2 ，其中 $E(C_1) < E(C_2)$ 而 $D(C_1) > D(C_2)$ ，且在大多数情况下 $T_\alpha(C_1) > T_\alpha(C_2)$ 。人们的选择偏好差异将导致了他们进行不同的选择。有的人倾向于冒险，选择路径 C_1 ，而有的人倾向于保守，选择路径 C_2 而有的人介于两者之间。显然上述构造的优化准则只是其中一种准则而已。对于不同选择偏好的人，适宜构造不同的优化准则。

我们以概率保证值来衡量倾向冒险的程度，当选择不同的 α 时，选取的路径就可能不一样。 α 取得比较大的人群是倾向保守决策的人群， α 取得比较小的人群是倾向危险决策的人群。

1) 边分位数法 由于每条路上的通行时间是一随机变量而非确定的值，所以没有办法直接运用加权有向图中的最短路径算法。下面构造一个确定的权值以利用最短路算法。由概率论知识，我们可以求出每条边的通行时间 $t_{i,j}$ 在某个概率保证条件下的分位数。设 $t_{i,j}$ 在概率为 $1-\alpha$ 下的分位数为 $t_{i,j}^*$ ，则不同的路径的通行时间已经由随机变量转化为一个确定的值，以此值作为每条边的权值，运用图论中的最短路算法即可求解该概率意义下的最短路问题。则 $\min t = t^* = \min(t_{i,j}^*)$ 就是概率意义下的最短

通行时间。

而对应的路径就是最段时间路径。

- 2) 路径分位数法 上述边分位数法将每条路上的随机的通行时间转化为了一个确定的分位数,从而将问题转化为了图论中经典的最短路问题。但是上述处理过程中一个缺点就是:本来整个通行时间 t 是其路径通过的边上的所有通行时间的和。所以 t 也是一个随机变量。分段求分位数再累加得到的值并非 t_k 在该概率条件下的分位数。

假设总共有 m 条互不相同的可行路径,则每条路径的通行时间 t_k 也同样是服从均值为 μ_k 均方差为 σ_k 的正态分布。假设优化准则为 t_k 在某个概率保证条件下的分位数。设 t_k 在概率为 $1-\alpha$ 下的分位数为 t_k^* ,则不同的路径的通行时间已经由随机变量转化为了一个确定的值,以此为目标函数,求出 t_k^* ($k=1,2,\dots,m$)的最小值 $\min t_k^*$,则 $\min t_k^*$ 就是概率意义下的最短通行时间。而对应的路径就是最段时间路径。

一般情况下同样概率下路径分位数法的预测值与边分位数法的预测值不一样。对于既不保守又不冒进的人群选取 $\alpha=50$ 时,显然该分位数就是 t_k 的均值 μ_k 。从而路径分位数法与边分位数法的结果一致,这是一个特例。

4.1.2 优化算法 边分位数法中,将未知的问题转化为了经典的图论最短路算法,最著名的譬如 Dijkstra 算法等,具体方法从略。而路径分位数法中权重依旧是随机变量。为此,我们有必要计算出所有可行的路径中所有路径的通行时间 t_k ($k=1,2,\dots,m$)的均值和方差,然后再求每条路径在概率为 $1-\alpha$ 下的分位数为 t_k^*

寻找所有可行路径的算法采用生成树的方法。生成树的根节点为出发点第 i 节点,所有的叶子节点都为终点第 j 节点。从根节点到任意一个叶子节点的路径就为一条可行路径。如果有 m 个叶子节点,则有 m 个可行路径。

生成树的算法如下:当我们从 i 节点出发,搜索所有与其相连接的边,沿这些边并行前进到下一层节点。检查路径中的节点编号是否与已经经过的节点重复。如果重复,则去掉该路径。然后再搜索下一层节点的连接的边,再判断每条边是否与节点与前面已经经过的节点重复(重复表示该路形成了圈,显然一定不是最好的路径)依次类推。

当所有的路径都前进到了目的地 j 节点后,计算结束。生成了所有的可行路径。实际上,当程序完成了生成树的同时,已经累加出了每条路径上的通行时间的均值和方差。再分别由正态分布找出其对应的 $1-\alpha$ 下的分位数,再对这 m 个分位数采用冒泡法排序,找出了概率意义下的最优解、次优解。

4.2 时变相关随机交通网络模型

上述交通网络没有考虑时间的影响。实际上,道路交通各特征参数不仅仅是随机的,且其数字特征也随时间的变化发生显著的变化。同时,由于交通网络是个有机的整体,所以其各参数之间不可避免地存在相关性。所以上述定常统计模型在时间预测和最短时

间路径选择上在一些情况下不是十分合理。下面对各随机变量的数字特征对时间的依赖和各随机变量之间的相关性做个简单分析。

4.2.1 时变系统的时间域迭代算法 由于一般情况下交通的状况都是以一昼夜为周期的准周期变化过程，有高峰期和低峰期。且由人们的出行习惯决定了高峰、低峰在相对固定的时间内出现。所以可以将 24 小时分割成半个小时为单位的 48 个时间等份。在每个时间段 t_i 上分别统计各边的通行时间的均值和方差，将其表示为 $\mu_{i,j}(t_i)$ 、 $\sigma_{i,j}(t_i)$ 。如果 $\mu_{i,j}(t_i)$ 、 $\sigma_{i,j}(t_i)$ 随时间的变化比较快（这在高峰期的形成和消退期可能表现得尤为显著），再考虑 t_0 时刻从节点 i 到节点 j 的时间预测问题：设途中经过 n 个边，编号暂时记作 l_k ，需要的时间为：

$$\Delta T(n_k) = \sum_{k=1}^{n_k} T_k \quad (34)$$

显然，考虑到系统的时变特性， T_k 依赖于时间 t 。当沿路径到达第 n_k 个边时，各边的数字特征已经变成了 $\mu_{i,j}(t_0 + \Delta t)$ 、 $\sigma_{i,j}(t_0 + \Delta t)$ 。所以必须对每个时间步骤更新系统的参数。现以每个边的通行时间为时间演化的步长，需要的时间为：

$$\bar{T}_n = \sum_{k=1}^n T_k(t_0 + \bar{T}_{k-1}) \quad (35)$$

其中 $T_k(t_0 + \Delta \bar{T}_{k-1})$ 表示该路径中经过的第 k 个边在 $t_0 + \Delta \bar{T}_{k-1}$ 时刻的通行时间。 \bar{T}_{k-1} 表示从 t_0 时刻起走过 $k-1$ 个边所需要的时间——即从出发至到达边 l_k 的入口所需要的时间， \bar{T}_n 表示从 t_0 时刻起走过 n 个边到达目的地 j 节点时所需要的时间。

显然，上面公式是迭代公式，且式中的数据都是随机变量，不可以直接迭代。而是在一定的概率条件下取值进行演化。首先由统计数据得到出 t_0 时刻，途径的第一条边的通行时间的均值和方差，计算出在一定的概率保证下到达下条边入口节点所需要花费的时间。由此时间计算时间由 t_0 演化到的新的时刻 $t_0 + \Delta t$ 。再由得到的这个新的时刻，根据统计数据得到此时刻第二条边的通行时间的均值和方差。依次计算出在该概率保证下的通过各边所需要的通行时间，并将得到的时间累加到演化时间上去，不断地更新系统的演化时间将系统向下演化，直到到达节点 j 为止。

4.2.2 由流量的相关性构造协方差矩阵 从前面一维模型的分析中和实际的监测数据中，我们都可以看到，由于当车流量过饱和的时候，会在一些交通不是很畅通的位置形成堵塞从而增加了车的密度，从而降低了车速。所以在车流量比较大的前提下，假定各边上的通行时间与其流量成线性关系。

在交通网络中，由于各节点上车流量守恒，设在 i 节点有 n 条边相连，每个边的流量为 $q_{i,j}$ 则必有：

$$\sum_{j=1}^n q_{i,j} \quad (36)$$

即一个边上的流量为其他边上的流量的线性组合。所以可以推断出：相临（拥有共同节点）的边上的流量是线性相关的；不相临的边之间尽管有影响，但是该影响由于流量在各个节点的进一步分散而减弱，故忽略不考虑。由于前面假设通行时间与流量成线性关系。故相临节点上的通行时间成线性相关关系，非相临的边依旧忽略不考虑。

进一步，构造出边 $l_{i,j}$ 与边 $l_{k,l}$ 之间的相关关系，假定其相关系数为：

$$\rho(ij,kl) = 0 \quad (i,j,k,l \text{ 各不相同, 表示不相临边}) \quad (37)$$

$$\rho(ij,il) = \gamma \frac{T_{i,j}}{\sum T_i} \frac{T_{i,l}}{\sum T_i} \quad (i \text{ 节点相临的两条边}) \quad (38)$$

其中 γ 为一合适的调节系数。显然，上述构造满足 $-1 \leq \rho(ij,kl) \leq 1$ 和对称性，满足相关系数的特征。同时 $\rho(ij,kl)$ 总为正数。其物理意义也是明显的：一般情况下，当一个交叉路口发生交通堵塞的时候，车辆往往想绕行，从而导致连接该节点的其他边上的车流量增加即通行时间增加。而一条路车辆比较少的时候，必然将其同一个节点上的其他边上的部分车流转移过去，从而减少其他路上的通行时间。所以相临边上的通行时间总是正相关的。

由概率论知识有

$$\text{Cov}(x,y) = \rho_{xy} \sqrt{DX} \sqrt{DY} \quad (39)$$

由式 (39) 可以分别计算出 $\text{Cov}(ij,kl)$ ，从而构造出协方差矩阵。

4.2.3 由历史监测数据直接按定义计算协方差矩阵

实际的交通状况是很复杂的，也并非完全随机的。上述构造的相关系数一定程度上地描述了边之间通行时间的相互影响、相互制约的关系。但是可能并不能很好地吻合实际情况。文献[7][8]也分别给出了类似的构造方法，但是共同的缺点就是：虽然构造算法有一定的物理意义，但是认为引入的假设太多，从而可能导致算法的结果与实际偏差过大。为此，我们寻找另外的直接由数据出发直接计算协方差系数的方法。

由概率论知识有：

$$\text{Cov}(x,y) = E(XY) - E(X)E(Y) \quad (40)$$

我们首先将实际测量数据按照时间分段，根据大量的历史数据资料在每个时间段上分别求出各边 $l_{i,j}$ 的通行时间的均值与方差，同时再计算出相应的 $E(l_{i,j}l_{k,l})$ 。则由公式 (40) 可以直接按照定义计算出 $\text{Cov}(l_{ij},l_{kl})$ 。

直接按照定义计算的方法虽然很直接，不用担心误差，但是当数据量比较大的时候， $E(l_{i,j}l_{k,l})$ 的计算过程十分耗时。所以有必要将构造算法与直接计算结合起来扬长避短。首先我们可以由具备一定物理意义的假设进行构造，然后用按照定义计算的结果与其比较并进行修正。这样用定义计算值作为标准，在反复实践的基础上可能构造出计算量小且精度可以接受的协方差矩阵的构造算法。

4.2.3 时变相关随机网络的最短时间路径选择算法 上述非时变的，各变量相互独立的交通网络简化成了一个含有随机权重的网络的最短路问题。由于走不同的路径的时间为随机变量，最后通过概率方法来确定概率意义下的时间最小路径。

由概率论，有：

$$D(\sum_{i=1}^n X_i) = \sum_{i=1}^n D(X_i) + 2 \sum_{1 \leq i < j \leq n} Cov(X_i, X_j) \quad (41)$$

$$E(\sum_{i=1}^n X_i) = \sum_{i=1}^n E(X_i) \quad (42)$$

由公式（41）（42），根据构造的或者按照定义计算得到的各边在某时段的通行时间的均值、方差与协方差矩阵就可以计算出任意一条路径的通行时间 t_k 的均值与方差。考虑系统的时变特性，按照公式（34）（35）描述的迭代算法，不断迭代到新的时刻并获取新时刻的系统参数。最终可以得到所有可行路径的在某给定的概率条件下的通行时间。其中最小的那个就是最短通行时间，其对应的路径就是最优路径。

五、算例

在上述模型中，各边通行时间的均值和方差需要由大量的数据统计得出。赛题中给出了两个假设：

- 1) 假设各边的通行时间的均值与边的长度成正比： $\mu_{ij} = K_{ij} S_{ij}$
- 2) 假设各边的通行时间的方差与长度的 $2/3$ 次方的倒数以及每个节点所连接的边数的乘积成正比——即与其两端节点的度成正比： $\sigma_{ij} = R_{ij} p_i p_j \left(\frac{1}{S_{ij}^{2/3}} \right)$

然后根据图的颜色大致估计出各路段的平均速度由此确定大致的 K_{ij} ，由这两个假设可以计算出每边通行时间的均值和方差。将数据代入上述稳态独立随机网络模型，运用分步算法，取 $\alpha=0.05$ ，可以得到最优路径为：

14 → 3 → 6 → 9 → 8 → 7 → 11 → 14

14 → 11 → 7 → 8 → 9 → 6 → 3

（附注：最优路径与其对应的最短时间强烈依赖具体参数，由于赛题中没有明确给出各系数的取值，所以该结果可能与用其他参数求的结果不同）

六、结论与展望

本文在对交通微观模型与随机交通网络的独立研究基础上，得出了如下重要结论：

- 1) 建立了交通系统的微观物理模型，并由微观模型推导出了交通系统的宏观特性。同时通过理论分析和 TransGuide 实际测量数据验证了模型的有效性和正确性。
- 2) 道路的极限运输能力由反应时间决定而与流量、速度等无关。

- 3) 交通流随密度的不同规律也不同,可以分为轻载、满载和过载三种情况。轻载时平均速度为常数 v^* ,密度与流量成正比,流量成泊松分布;满载时速度与密度成反比关系;过载时间过长将使得交通系统逐渐转入堵车状态,运输能力并不下降但是道路的通行时间大大增加。
- 4) 堵车时的流量依旧为极限运输流量而不是零,对应的平均速度为最低值 v_* ,密度为极限密度 ρ^* 而有效密度 ρ_e 趋向无穷大。
- 5) 最优路径的优化目标函数需要考虑主体选择的差异性。通过分位数统一了目标函数,将随机网络转化成了图论中的最短路问题。
- 6) 分析了时变随机交通网络的时变特性,并给出了求解具有显著时变特性交通系统的长时间尺度时间预测问题的迭代算法。
- 7) 分析了随机交通网络中各边的特征参数的相关性,并构造了协方差矩阵来描述相关性。

由于时间仓促,本文的模型还需要进一步完善。主要有以下工作要做:从 transguide 系统的网站获得大量的实际监测数据,进一步统计分析随机交通网络中的各变量间的关系。并考虑到由于司机的反馈,系统将不仅仅是以统计规律变化,还将体现出作为大系统的负反馈特性。故需要将司机的主体行为进一步加入模型。

七、参考文献

- 1、TransGuide 系统——ITS 系统示范项目 鲁忠武
<http://tte.whut.edu.cn/admin/show.php?dno=179>
- 2、城市交通流路段行程时间预测模型.杨昊.北方交通大学学报.第 25 卷.第 2 期.
- 3、短时交通流智能预测方法的研究.王伟平.山东交通学院学报.第 12 卷.第 2 期.
- 4、基于 ARIMA 模型的短时交通流实时自适应预测.韩 超.系统仿学报.16卷.第7期
- 5、Dailey D J. travel time estimation using cross-correlation technique. Transpn res. 1993 27B(2) 97~107
- 6、Essentially Best Routes in Dynamic and Stochastic **Transportation Network**, Rachel R. He¹, Alain L. Kornhauser², Bin Ran³, University of Wisconsin.
- 7、Estimation of Time-dependent O-D Demand and Route **Choice from Link Flows** Rachel R. He¹, Alain L. Kornhauser², Bin Ran³.
- 8、A General Purpose Methodology for Link Travel Time Estimation Using Multiple Point Detection of Traffic. **Cristián E. Cortés**. Department of Civil Engineering and Institute of Transportation Studies